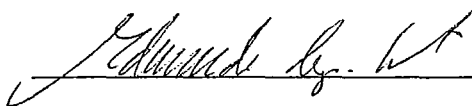


EXPERIMENTOS COM O SERVIDOR RIO EM UM AMBIENTE
DISTRIBUÍDO E HETEROGÊNEO

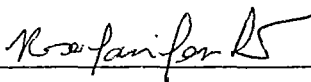
Vinicius Martins Botelho

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO
DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

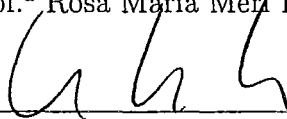
Aprovada por:



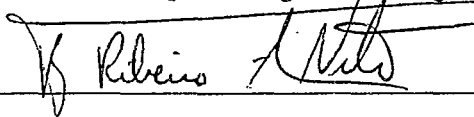
Prof. Edmundo Albuquerque de Souza e Silva, Ph.D.



Prof.ª Rosa Maria Merl Leão, Dr.



Prof. Paulo Henrique de Aguiar Rodrigues, Ph. D.



Prof. Berthier Ribeiro de Araújo Neto, Ph. D.

RIO DE JANEIRO, RJ - BRASIL

JULHO DE 2006

BOTELHO, VINICIUS MARTINS

Experimentos com o Servidor RIO em um ambiente distribuído e heterogêneo [Rio de Janeiro] 2006

XVII, 100 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 2006)

Dissertação - Universidade Federal do Rio de Janeiro, COPPE

1. Balanceamento de Carga
2. Vídeo sob Demanda
3. Redes Gigabit
4. Diversidade de Caminhos

I. COPPE/UFRJ II. Título (série)

Aos meus pais, pedra angular de minha vida.

Agradecimentos

Em primeiro lugar, devo agradecer a Deus, por ter pensado em mim desde a eternidade, ter me criado e me sustentado até hoje. Agradeço também aos meus pais, que são co-responsáveis, em parceria com Deus, pela minha existência. Além disso, é o seu amor por mim que me faz capaz de compreender, ainda que palidamente, a infinidade do Amor divino. E também devo agradecer os seus incentivos constantes e ajuda sem limite que me prestaram durante toda a minha vida.

Quero agradecer muitíssimo aos professores Edmundo e Rosa, que sempre confiaram e acreditaram em mim, mesmo nos momentos em que ninguém o faria. Espero que este trabalho seja minha retribuição pela confiança depositada em mim, desde 2001, quando fui aluno do professor Edmundo em TP.

Um agradecimento muito especial está reservado para a secretária mais querida da UFRJ: a Carol. Ela foi calma quando eu era intempestivo, ria quando eu chorava e ouvia quando eu precisava desabafar.

Tenho que agradecer a todos meus amigos e amigas que, de alguma forma, contribuíram para a conclusão deste trabalho: Primeiramente devo agradecer a Juliana Cunha, pelo incentivo, paciência e correção do *abstract*. Um agradecimento muito importante deve ser feito ao Bernardo Netto, tamanha foi sua ajuda na implementação no RIO e outro ao Flávio, por me ajudar a destrinchar o bisonho comportamento do *kernel* do Linux. Quero agradecer em especial os administradores do LAND, Boechat, Hugo e Carolzinha, por sempre concederem os *sudos* que eu pedia prontamente. E a todos do LAND: GD, Guto, Fabianne, Ana Paula, Allyson, Bene, Fernando, Sadoc, Diana, Isabela, Watanabe e demais, pelo companheirismo e prontidão no auxílio!

Agradeço aos "garotos e garota" que trabalharam comigo durante o Projeto Giga/RNP: Bruno, Marcello e Ariadne. Este trabalho foi uma consequência do

Projeto Giga/RNP: agradeço a RNP e a FINEP pelo financiamento e pela realização do projeto agradeço a COPPE/UFRJ, DCC/UFF, DCC/UFMG e ao Canal Saúde/FIOCRUZ. Deste projeto veio parte dos recursos de minha bolsa e agradeço ao FAPERJ, que fez a complementação na etapa final. Também agradeço a Fundação Coppetec, que viabilizou financeiramente meus estudos desde a metade de minha graduação e foi a forma pela qual recebi a bolsa do Projeto Giga/RNP.

Por fim, mas não menos importante, agradeço ao Bernardo Miranda, que me deu uma carona vital para cumprir um prazo. Também agradeço aos amigos da UFF, professora Anna Dolejsi e Sandoval, aos amigos da UFMG, professora Jussara e Alex Borges e aos amigos do Canal Saúde, Angélica, Odir, Mirela e Neide.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

EXPERIMENTOS COM O SERVIDOR RIO EM UM AMBIENTE
DISTRIBUÍDO E HETEROGÊNEO

Vinicius Martins Botelho

Julho/2006

Orientador: Edmundo Albuquerque de Souza e Silva

Programa: Engenharia de Sistemas e Computação

Nesta dissertação, realizamos uma série de experimentos no servidor multimídia RIO em um ambiente real que incluiu a Internet e uma rede paralela com velocidades de gigabits por segundo. Assim, foi possível verificar possíveis gargalos de desempenho e estudar maneiras de contorná-los. Como o ambiente de experimentação inclui uma rede com caminhos heterogêneos, com taxas de perda e tempo de propagação claramente distintos, também foi possível traçar um paralelo com um problema conhecido na literatura como diversidade de caminhos. Utilizamos métricas para avaliar o desempenho do sistema, como a taxa de *goodput* para o cliente. Assim, avaliamos qual o impacto, para a qualidade de serviço (QoS), da adição de nós extras de armazenamento e comparamos os resultados da adição em cada rede. Também investigamos o impacto da replicação dos vídeos na QoS do sistema, além do aumento de confiabilidade. Concluímos que a adição de um nó de armazenamento na rede contribui positivamente para a capacidade total do sistema, mesmo que esta introduza alta latência e taxa de perda.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

EXPERIMENTS WITH THE RIO SERVER IN A DISTRIBUTED AND
HETEROGENEOUS ENVIRONMENT

Vinicius Martins Botelho

July/2006

Advisor: Edmundo Albuquerque de Souza e Silva

Department: Systems Engineering and Computer Science

In this dissertation, we have done experiments with the RIO multimedia server in a real environment that includes the Internet and an additional network with gigabit per second bandwidth. We investigated the potential performance bottlenecks and we studied ways to circumvent them. As the experimentation environment had heterogeneous paths, with clearly distinct loss rates and propagation delays, it was also possible to relate our scenario with the problem known in the literature as path diversity. We used known metrics to measure the system's performance, such as the client's goodput rate. We measured the impact on the Quality of Service (QoS) of adding extra storage nodes both in the gigabit network and in the Internet. We also investigated the possible performance improvements of block replication in the system's QoS. We conclude that the addition of a storage node to the system can have a positive impact to the overall system performance, even if this node is connected to the server through a network with higher packet loss rate and delay.

Lista de Acrônimos

VoD :	<i>Video on Demand;</i>
DVoD :	<i>Distributed Video on Demand;</i>
P2PVoD :	<i>Peer-to-Peer Video on Demand;</i>
RIO :	<i>Randomized I/O Multimedia Server;</i>
Mbps :	<i>Megabits per second;</i>
KB :	<i>Kilobytes;</i>

Sumário

Resumo	vi
Abstract	vii
Lista de Acrônimos	viii
Lista de Figuras	xiii
Lista de Tabelas	xvii
1 Introdução e Motivação	1
1.1 Principais objetivos	2
1.2 Principais contribuições	2
1.3 Por que estudar vídeo sob demanda?	3
1.4 Interativo x Ao Vivo: Problemas inerentes a cada tipo de serviço	5
1.5 Organização do trabalho	7
2 Principais Arquiteturas para <i>Streaming</i> Multimídia	8
2.1 Quanto a função dos agentes	8
2.1.1 Arquiteturas VoD	8
2.1.2 Arquiteturas DVoD	10
Proxies: descentralização parcial	10
CDNs: marketing ou ganho real?	12
2.1.3 Arquiteturas Peer-to-Peer	13
Definição de uma arquitetura <i>peer-to-peer</i>	14
Definição de uma arquitetura com o uso de computação <i>Grid</i>	14

	<i>Grid x Peer-to-Peer</i>	15
	Arquiteturas P2P quanto à organização e função dos <i>peers</i>	15
	Arquiteturas P2P quanto à distribuição dos dados	18
	Problemas específicos às arquiteturas P2P	19
2.2	Quanto à forma de <i>streaming</i>	21
2.3	Quanto à forma de alocação dos vídeos	23
2.3.1	Formas de dispersão dos blocos nos discos	23
2.3.2	Formas de replicação dos vídeos	25
2.4	Quanto ao caminho	26
3	Arquitetura do RIO	29
3.1	Origem e evolução do RIO	29
3.2	<i>Framework</i> das entidades do RIO	31
3.2.1	O Objeto Servidor	32
	Gerenciador de Sessão (<i>SessionManager</i>)	32
	Gerenciador de Eventos (<i>EventManager</i>)	33
	Gerenciador de Fluxos (<i>StreamManager</i>)	33
	Gerenciador de Objetos (<i>ObjectManager</i>)	34
	Gerenciador de Disco (<i>DiskManager</i>)	35
	Roteador (<i>Router</i>)	36
3.2.2	Objeto de armazenamento	36
	Interface com o Roteador (<i>RouterInterface</i>)	36
	Dispositivo de Armazenamento (<i>StorageDevice</i>)	37
	Gerenciador de Armazenamento (<i>StorageManager</i>)	37
	Interface com o Cliente (<i>ClientInterface</i>)	37
3.2.3	O Objeto Cliente	38
3.3	Controle de admissão	39
3.4	Policimento de pedidos	40
3.5	Contribuições deste trabalho ao Servidor RIO	41
3.5.1	Estratégia de replicação	41
3.5.2	Impressão de informações pelo cliente	42

4	Arquitetura da Rede Giga/RNP, Experimentos Realizados e Resultados	44
4.1	Objetivo dos Experimentos	44
4.2	Arquitetura da Rede Giga	45
4.2.1	Configuração dos Equipamentos Utilizados	48
4.3	Metodologia dos Experimentos	48
4.3.1	Duração do experimento	51
4.4	Primeiro Experimento: RIO e a rede Giga	52
4.4.1	Descrição dos objetivos	52
4.4.2	Descrição dos resultados esperados	52
4.4.3	Resultados e Análise	53
4.5	Segundo Experimento: Redundância no Servidor	55
4.5.1	Descrição dos objetivos	55
4.5.2	Descrição dos resultados esperados	56
4.5.3	Resultados e Análise	56
4.6	Terceiro Experimento: Carga de 1 nó de armazenamento com interface gigabit	59
4.6.1	Descrição dos objetivos	59
	Cálculo do <i>buffer</i> de um <i>switch</i>	59
4.6.2	Descrição dos objetivos: Continuação	61
4.6.3	Descrição dos resultados esperados	61
4.6.4	Resultados e Análise	62
4.7	Quarto Experimento: Análise dos caminhos entre os nós de armazenamento e os clientes	64
4.7.1	Descrição dos objetivos	64
4.7.2	Descrição dos resultados esperados	65
4.7.3	Resultados e Análise	66
4.8	Quinto Experimento: RIO na UFMG	67
4.8.1	Descrição dos objetivos	67
4.8.2	Descrição dos resultados esperados	67
4.8.3	Resultados e Análise	67

4.9	Sexto Experimento: RIO na UFMG com redundância	70
4.9.1	Descrição dos objetivos	70
4.9.2	Descrição dos resultados esperados	70
4.9.3	Resultados e Análise	71
4.10	Comparação entre o RIO na rede Giga e com o nó extra na UFMG	73
4.10.1	Descrição dos objetivos	73
4.10.2	Descrição dos resultados esperados	73
4.10.3	Resultados e Análise	73
4.11	Considerações sobre a escalabilidade do servidor RIO	79
5	Conclusão e Trabalhos Futuros	82
5.1	Trabalhos Futuros	83
	Referências Bibliográficas	85
A	Detalhes das Implementações Realizadas	93
A.1	Algoritmo de Cópia com Redundância	93
A.2	Melhorias no código	95
A.3	Formato do log gerado pelo cliente	97

Lista de Figuras

2.1	Exemplo de uma rede com um servidor multimídia único.	9
2.2	Exemplo de uma rede com proxies perto dos grupos de clientes. . . .	10
2.3	Exemplo de uma rede com proxies hierárquicos.	11
2.4	Exemplo de rede com a distribuição de conteúdo multimídia feito por uma CDN.	13
2.5	Exemplo da partição do fluxo de acordo com a capacidade dos nós. Da taxa total R necessária, o nó P2 tem mais banda disponível, logo ele é responsável por transmitir $R/2$ enquanto que os dois outros transmitem $R/4$	16
2.6	Exemplo de uma rede P2P que inicializa com um nó de inicialização (<i>bootstrap</i>) e que é organizada em dois níveis, de acordo com a capacidade dos nós. Os círculos preto maiores indicam nós com maior capacidade.	20
2.7	Exemplo de uma execução de um algoritmo de flood como o encontrado no Gnutella. Em (a) o nó inquisidor manda para todos seus vizinhos a sua requisição. Em (b) os nós vizinhos repassam a requisição para seus vizinhos, uma vez que eles não possuem o arquivo procurado. Em (c) é encontrado o arquivo e a resposta é enviada. Em (d) a resposta é repassada ao nó inquisidor original.	22
3.1	Arquitetura do RIO.	32
3.2	Diagrama de classes dos componentes do RIO.	33
3.3	Arquivo de metadados de um objeto do RIO com três cópias.	35
3.4	Fluxo lógico do <i>Router</i>	37

3.5	Diagrama de Classes do nó de armazenamento.	38
3.6	O fluxo da troca de informações entre os agentes do sistema durante o processo de cópia de um novo objeto.	42
4.1	Representação simplificada da VPN da rede Giga.	45
4.2	Representação simplificada da arquitetura da rede no laboratório da UFRJ.	46
4.3	Representação simplificada da arquitetura de todo o ambiente de testes entre as instituições participantes dos experimentos.	47
4.4	Gráfico do número de clientes x <i>Goodput</i> com o RIO somente na rede Giga, de 1 a 4 nós de armazenamento, sem réplicas.	53
4.5	Gráfico do número de clientes x porcentagem de fragmentos atrasados com o RIO somente na rede Giga, de 1 a 4 nós de armazenamento, sem réplicas.	54
4.6	Gráfico do número de clientes x porcentagem de fragmentos perdidos com o RIO somente na rede Giga, de 1 a 4 nós de armazenamento, sem réplicas.	54
4.7	Gráfico do número de clientes x <i>Goodput</i> com o RIO somente na rede Giga, de 2 a 4 nós de armazenamento, com réplicas.	56
4.8	Gráfico do número de clientes x porcentagem de fragmentos atrasados com o RIO somente na rede Giga, de 2 a 4 nós de armazenamento, com réplicas.	57
4.9	Gráfico do número de clientes x porcentagem de fragmentos perdidos com o RIO somente na rede Giga, de 2 a 4 nós de armazenamento, com réplicas.	57
4.10	Gráfico do número de clientes x <i>Goodput</i> com o RIO somente na rede Giga, comparando 4 nós de armazenamento com e sem réplicas.	58
4.11	Gráfico do número de clientes x <i>Goodput</i> com o RIO em uma interface de rede 1000, 1 nó de armazenamento.	62
4.12	Gráfico do número de clientes x porcentagem de fragmentos atrasados com o RIO em uma interface de rede 1000, 1 nó de armazenamento.	63

- 4.13 Gráfico do número de clientes x porcentagem de fragmentos perdidos com o RIO em uma interface de rede 1000, 1 nó de armazenamento. . . 63
- 4.14 Gráfico do número de clientes x goodput com o RIO em uma interface de rede 1000, 1 nó de armazenamento, e o RIO na rede Giga, com 4 nós de armazenamento e replicação. 64
- 4.15 Gráfico do número de clientes x *Goodput* com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, sem réplicas. 68
- 4.16 Gráfico do número de clientes x porcentagem de fragmentos atrasados com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, sem réplicas. 69
- 4.17 Gráfico do número de clientes x porcentagem de fragmentos perdidos com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, sem réplicas. 69
- 4.18 Gráfico do número de clientes com *goodput* de até 96% x número de nós de armazenamento somente na rede Giga em comparação com a mesma configuração adicionando-se o nó extra na UFMG, sem réplicas. 70
- 4.19 Gráfico do número de clientes x *Goodput* com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, com replicações. 71
- 4.20 Gráfico do número de clientes x porcentagem de fragmentos atrasados com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, com replicações. 72
- 4.21 Gráfico do número de clientes x porcentagem de fragmentos perdidos com o RIO na rede Giga e sempre com 1 nó na UFMG, de 2 a 5 nós de armazenamento, com replicações. 72
- 4.22 Gráfico do número de clientes x *Goodput* com o RIO com 1 nó na rede Giga em comparação com 1 nó na rede Giga e 1 nó na UFMG, sem replicações. 75

4.23	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 2 nós na rede Giga em comparação com 1 nó na rede Giga e 1 nó na UFMG, sem replicações.	75
4.24	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 2 nós na rede Giga em comparação com 2 nós na rede Giga e 1 nó na UFMG, sem replicações.	76
4.25	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 3 nós na rede Giga em comparação com 2 nós na rede Giga e 1 nó na UFMG, sem replicações.	76
4.26	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 3 nós na rede Giga em comparação com 3 nós na rede Giga e 1 nó na UFMG, sem replicações.	77
4.27	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 4 nós na rede Giga em comparação com 3 nós na rede Giga e 1 nó na UFMG, sem replicações.	77
4.28	Gráfico do número de clientes x <i>Goodput</i> com o RIO com 4 nós na rede Giga em comparação com 4 nós na rede Giga e 1 nó na UFMG, sem replicações.	78
4.29	Gráfico da taxa de ocupação do processador x tempo de simulação, com o RIO com 4 nós na rede Giga e 200 clientes.	80
4.30	Gráfico da taxa de ocupação do processador x tempo de simulação, com o RIO com 4 nós na rede Giga e 300 clientes.	80
4.31	Gráfico da taxa de ocupação da memória RAM x tempo de simulação, com o RIO com 4 nós na rede Giga e 200 clientes.	81
4.32	Gráfico da taxa de ocupação da memória RAM x tempo de simulação, com o RIO com 4 nós na rede Giga e 300 clientes.	81

Lista de Tabelas

1.1	Diferenças entre um serviço de vídeo sob demanda (VoD) e um serviço de transmissão ao vivo.	6
2.1	Diferenças básicas entre arquiteturas <i>peer-to-peer</i> (P2P) e <i>Grid</i>	15
2.2	Classificação das propostas de arquiteturas para streaming multimídia.	28
4.1	Comparação entre os RTTs através da aplicação que foram medidos na rede Giga e na Internet entre a UFRJ e UFMG, em milissegundos.	66

